WHAT IS CLAIMED IS:

1. A speech processing apparatus comprising:

generation means for generating a pseudo acoustic echo signal based on a current impulse response simulating an acoustic echo transfer path and on a source signal;

sub B1

supply means for holding the current impulse response and supplying the current impulse response to said generation means;

elimination means for subtracting said pseudo acoustic echo signal from a microphone input signal to remove an acoustic echo component and thereby generate an acoustic echo-canceled signal;

update means for continually updating the impulse response by using said source signal, said acoustic echo-canceled signal and the current impulse response held by said supply means and for supplying the updated impulse response to said supply means;

decision means for checking, in each frame, whether or not a voice is included in the microphone input signal, by using time domain information and frequency domain information of said acoustic echo-canceled signal;

storage means for storing one or more impulse

responses; and

sub B1 cont

~~control means for, in a frame for which the result of decision made by said decision means is negative, storing in said storage means the current impulse response held by said supply means and, in a frame for which the result of decision is positive, retrieving one of the impulse responses stored in said storage means and supplying it to said supply means.~~

2. A speech processing apparatus as claimed in claim 1, wherein said acoustic echo-canceled signal is used for speech recognition.

3. A speech processing apparatus as claimed in claim 2, further comprising:

means for determining a spectrum for each frame by performing the Fourier transform on said acoustic echo-canceled signal;

means for successively determining a spectrum mean for each frame based on the spectrum obtained; and

means for successively subtracting the spectrum mean from the spectrum calculated for each frame from said acoustic echo-canceled signal to remove additive noise of an unknown source.

4. A speech processing apparatus as claimed in claim 2, further comprising:

means for determining a spectrum for each frame by performing the Fourier transform on said acoustic echo-canceled signal;

means for successively determining a spectrum mean for each frame based on the spectrum obtained;

means for successively subtracting the spectrum mean from the spectrum calculated for each frame from said acoustic echo-canceled signal;

means for determining a cepstrum from the spectrum, the spectrum being removed of the additive noise of an unknown source by said subtraction means;

means for determining for each talker a cepstrum mean of a speech frame and a cepstrum mean of a non-speech frame, separately, from the cepstrums obtained; and

means for subtracting the cepstrum mean of the speech frame of each talker from the cepstrum of the speech frame of the talker and for subtracting the cepstrum mean of the non-speech frame of each talker from the cepstrum of the non-

speech frame of the talker to correct multiplicative distortions that are dependent on microphone characteristics and spatial transfer characteristics from the mouth of the talker to the microphone.

5. A speech processing apparatus as claimed in claim 2, further comprising:

means for determining a spectrum for each frame by performing the Fourier transform on said acoustic echo-canceled signal;

means for determining a cepstrum from the spectrum obtained; means for determining for each talker a cepstrum mean of a speech frame and a cepstrum mean of a non-speech frame, separately, from the cepstrums obtained; and

means for subtracting the cepstrum mean of the speech frame of each talker from the cepstrum of the speech frame of the talker and for subtracting the cepstrum mean of the non-speech frame of each talker from the cepstrum of the non-speech frame of the talker to correct multiplicative distortions that are dependent on microphone characteristics and spatial transfer characteristics from the mouth of the talker to the microphone.

6. A speech processing apparatus comprising:

means for determining a spectrum for each frame by the Fourier transform;

means for determining a cepstrum from the spectrum obtained;

means for determining for each talker a cepstrum mean of a speech frame and a cepstrum mean of a non-speech frame, separately, from the cepstrums obtained; and

means for subtracting the cepstrum mean of the speech frame of each talker from the cepstrum of the speech frame of the talker and for subtracting the cepstrum mean of the non-speech frame of each talker from the cepstrum of the non-speech frame of the talker to correct multiplicative distortions that are dependent on microphone characteristics and spatial transfer characteristics from the mouth of the talker to the microphone.

7. A speech processing method comprising:

a generation step for generating a pseudo acoustic echo signal based on a current impulse response simulating an acoustic echo transfer path and on a source signal;

Sub E4

a supply step for holding the current impulse response and supplying the current impulse response to said generation step;

an elimination step for subtracting said pseudo acoustic echo signal from a microphone input signal to remove an acoustic echo component and thereby generate an acoustic echo-canceled signal;

an update step for continually updating the impulse response by using said source signal, said acoustic echo-canceled signal and the current impulse response held by the supply step and for supplying the updated impulse response to said supply step;

a decision step for checking, in each frame, whether or not a voice is included in the microphone input signal, by using time domain information and frequency domain information of said acoustic echo-canceled signal;

a storage step for storing one or more impulse responses; and

a control step for, in a frame for which the result of decision made by said decision step is negative, storing in said storage step the current impulse response held by the supply means and, in a frame for which the result of decision is

positive, retrieving one of the impulse responses stored in said storage step and supplying it to said supply step.

8. A speech processing method as claimed in claim 7, wherein said acoustic echo-canceled signal is used for speech recognition.

9. A speech processing method as claimed in claim 8, further comprising:

a step for determining a spectrum for each frame by performing the Fourier transform on said acoustic echo-canceled signal;

a step for successively determining a spectrum mean for each frame based on the spectrum obtained; and a step for successively subtracting the spectrum mean from the spectrum calculated for each frame from said acoustic echo-canceled signal to remove additive noise of an unknown source.

10. A speech processing method as claimed in claim 8, further comprising:

a step for determining a spectrum for each frame by performing the Fourier transform on said acoustic echo-canceled signal;

a step for successively determining a

spectrum mean for each frame based on the spectrum obtained;

a step for successively subtracting the spectrum mean from the spectrum calculated for each frame from said acoustic echo-canceled signal to remove additive noise of an unknown source;

a step for determining a cepstrum from the spectrum removed of the additive noise;

a step for determining for each talker a cepstrum mean of a speech frame and a cepstrum mean of a non-speech frame, separately, from the cepstrums obtained; and

a step for subtracting the cepstrum mean of the speech frame of each talker from the cepstrum of the speech frame of the talker and for subtracting the cepstrum mean of the non-speech frame of each talker from the cepstrum of the non-speech frame of the talker to correct multiplicative distortions that are dependent on microphone characteristics and spatial transfer characteristics from the mouth of the talker to the microphone.

11. A speech processing method as claimed in claim 8, further comprising:

a step for determining a spectrum for each

frame by performing the Fourier transform on said acoustic echo-canceled signal;

a step for determining a cepstrum from the spectrum obtained; a step for determining for each talker a cepstrum mean of a speech frame and a cepstrum mean of a non-speech frame, separately, from the cepstrums obtained; and

a step for subtracting the cepstrum mean of the speech frame of each talker from the cepstrum of the speech frame of the talker and for subtracting the cepstrum mean of the non-speech frame of each talker from the cepstrum of the non-speech frame of the talker to correct multiplicative distortions that are dependent on microphone characteristics and spatial transfer characteristics from the mouth of the talker to the microphone.

12. A speech processing method comprising:

a step for determining a spectrum for each frame by the Fourier transform;

a step for determining a cepstrum from the spectrum obtained;

a step for determining for each talker a cepstrum mean of a speech frame and a cepstrum mean of a non-speech frame, separately, from the

cepstrums obtained; and

a step for subtracting the cepstrum mean of the speech frame of each talker from the cepstrum of the speech frame of the talker and for subtracting the cepstrum mean of the non-speech frame of each talker from the cepstrum of the non-speech frame of the talker to correct multiplicative distortions that are dependent on microphone characteristics and spatial transfer characteristics from the mouth of the talker to the microphone.